OPEN ACCESS

International
Journal
Of Modern Engineering Research (IJMER)

Multiexposure Image Fusion

Suhaina J[1], Bindu K. R[2]

[1]*Department of Electronics and Communication Engineering, FISAT/ Mahatma Gandhi University, India*
[2]*Department of Electronics and Communication Engineering, FISAT/ Mahatma Gandhi University, India*

Abstract: *Many applications such as robot navigation, defense, medical and remote sensing perform various processing tasks, which can be performed more easily when all objects in different images of the same scene are combined into a single fused image. In this paper, we propose a fast and effective method for image fusion. The proposed method derives the intensity based variations that is large and small scale, from the source images. In this approach, guided filtering is employed for this extraction. Gaussian and Laplacian pyramidal approach is then used to fuse the different layers obtained. Experimental results demonstrate that the proposed method can obtain better performance for fusion of all sets of images. The results clearly indicate the feasibility of the proposed approach.*

Keywords: *Gaussian Pyramid, Guided Filter, Image Fusion, Laplacian Pyramid, Multi-exposure images*

I. INTRODUCTION

Often a single sensor cannot produce a complete representation of a scene. Visible images provide spectral and spatial details, and if a target has the same color and spatial characteristics as its background, it cannot be distinguished from the background. Image fusion is the process of combining information from two or more images of a scene into a single composite image that is more informative and is more suitable for visual perception or computer processing. The objective in image fusion is to reduce uncertainty and minimize redundancy in the output while maximizing relevant information particular to an application or task. Given the same set of input images, different fused images may be created depending on the specific application and what is considered relevant information. There are several benefits in using image fusion: wider spatial and temporal coverage, decreased uncertainty, improved reliability, and increased robustness of system performance.

A large number of image fusion methods [1]–[4] have been proposed in literature. Among these methods, multiscale image fusion [2] and data-driven image fusion [3] are very successful methods. They focus on different data representations, e.g., multi-scale coefficients [5], [6], or data driven decomposition coefficients [3], [7] and different image fusion rules to guide the fusion of coefficients. The major advantage of these methods is that they can well preserve the details of different source images. However, these kinds of methods may produce brightness and color distortions since spatial consistency is not well considered in the fusion process. Spatial consistency means that if two adjacent pixels have similar brightness or color, they will tend to have similar weights. A popular spatial consistency based fusion approach is formulating an energy function, where the pixel saliencies are encoded in the function and edge aligned weights are enforced by regularization terms, e.g., a smoothness term. This energy function can be then minimized globally to obtain the desired weight maps. To make full use of spatial context, optimization based image fusion approaches, e.g., generalized random walks [8], and Markov random fields [9] based methods have been proposed. These methods focus on estimating spatially smooth and edge aligned weights by solving an energy function and then fusing the source images by weighted average of pixel values. However, optimization based methods have a common limitation, i.e., inefficiency, since they require multiple iterations to find the global optimal solution. Moreover, another drawback is that global optimization based methods may over-smooth the resulting weights, which is not good for fusion. An interesting alternative to optimization based method is guided image filtering [10]. The proposed method employs guided filtering for layer extraction. The extracted layers are then fused separately.

The remainder of this paper is organized as follows. In Section II, the guided image filtering algorithm is reviewed. Section III describes the proposed image fusion algorithm. The experimental results and discussions are presented in Section IV. Finally, Section V concludes the paper.

II. GUIDED IMAGE FILTERING

Guided filter is an image filter derived from a local linear model. It computes the filtering output by considering the content of a guidance image, which can be the input image itself or another different image. The guided filter can be used as an edge-preserving smoothing operator like the popular bilateral filter, but it has better behaviors near edges. The guided filter is also a more generic concept beyond smoothing: It can transfer the structures of the guidance image to the filtering output, enabling new filtering applications like dehazing and guided feathering. Moreover, the guided filter naturally has a fast and nonapproximate linear time algorithm, regardless of the kernel size and the intensity range. Currently, it is one of the fastest edge-preserving filters. Guided filter is both effective and efficient in a great variety of computer vision and computer graphics applications, including edge-aware smoothing, detail enhancement, HDR compression, image matting/feathering, dehazing, joint upsampling, etc.

The filtering output is locally a linear transform of the guidance image. On one hand, the guided filter has good edge-preserving smoothing properties like the bilateral filter, but it does not suffer from the gradient reversal artifacts. On the other hand, the guided filter can be used beyond smoothing: With the help of the guidance image, it can make the filtering output more structured and less smoothed than the input. Moreover, the guided filter naturally has an O(N) time (in the number of pixels N) nonapproximate algorithm for both gray-scale and high-dimensional images, regardless of the kernel size and the intensity range. Typically, the CPU implementation achieves 40 ms per mega-pixel performing gray-scale filtering. It has great potential in computer vision and graphics, given its simplicity, efficiency, and high-quality.

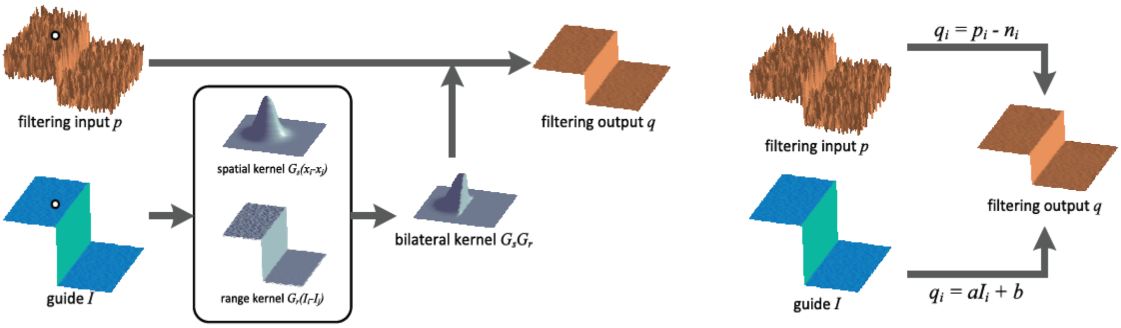

Fig. 2.1. Illustrations of the bilateral filtering process (left) and the guided filtering process (right)

2.1 Guided filter

A general linear translation-variant filtering process is defined, which involves a guidance image I, an filtering input image p, and an output image q. Both I and p are given beforehand according to the application, and they can be identical. The filtering output at a pixel i is expressed as a weighted average:

$$q_i = \sum_j W_{ij}(I)\, p_j \tag{1}$$

where i and j are pixel indexes. The filter kernel W_{ij} is a function of the guidance image I and independent of p. This filter is linear with respect to p. An example of such a filter is the joint bilateral filter (Fig. 2.1 (left)). The bilateral filtering kernel W_{bf} is given by :

$$W_{ij}^{bf}(I) = \frac{1}{K_i} exp\left(-\frac{\|x_i - x_j\|^2}{\sigma_s^2}\right) exp\left(-\frac{\|I_i - I_j\|^2}{\sigma_r^2}\right) \tag{2}$$

where x is the pixel coordinate and K_i is a normalizing parameter to ensure $\sum_j W_{ij}^{bf} = 1$. The parameters σ_s and σ_r adjust the sensitivity of the spatial similarity and the range (intensity/color) similarity, respectively. The joint bilateral filter degrades to the original bilateral filter when I and p are identical. The implicit weighted-average filters optimize a quadratic function and solve a linear system in this form:

$$Aq = p \tag{3}$$

where q and p are N-by-1 vectors concatenating $\{q_i\}$ and $\{p_i\}$, respectively, and A is an N-by-N matrix only depends on I. The solution to (3), i.e., $q = A^{-1}p$, has the same form as (1), with $W_{ij} = (A^{-1})_{ij}$.

The key assumption of the guided filter is a local linear model between the guidance I and the filtering output q. We assume that q is a linear transform of I in a window w_k centered at the pixel k:

$$q_i = a_k I_i + b_k, \forall i \epsilon w_k \tag{4}$$

where (a_k, b_k) are some linear coefficients assumed to be constant in w_k. We use a square window of radius r. This local linear model ensures that q has an edge only if I has an edge, because $\nabla q = a\nabla I$. This model has been proven useful in image super-resolution, image matting and dehazing.

To determine the linear coefficients {a_k, b_k}, we need constraints from the filtering input p. We model the output q as the input p subtracting some unwanted components n like noise/textures:

$$q_i = p_i - n_i \tag{5}$$

A solution that minimizes the difference between q and p while maintaining the linear model (4) is suggested. Specifically, the following cost function in the window w_k is minimized :

$$E(a_k, b_k) = \sum_{i\in w_k}((a_k I_i + b_k - p_i)^2 + \epsilon a_k^2) \tag{6}$$

Here, ϵ is a regularization parameter penalizing large a_k.
Equation (6) is the linear ridge regression model [11] and its solution is given by :

$$a_k = \frac{\frac{1}{|w|}\sum_{i\in w_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \epsilon} \tag{7}$$

$$b_k = \bar{p}_k - a_k \mu_k \tag{8}$$

Here, μ_k and σ_k^2 are mean and variance of I in w_k, |w| is the number of pixels in w_k, and $\bar{p}_k = \frac{1}{|w|}\sum_{i\in w_k} p_i$ is the mean of p in w_k. Having obtained the linear coefficients {a_k, b_k}, we can compute the filtering output q_i by (4). Fig.2.1 (right) shows an illustration of the guided filtering process.

However, a pixel i is involved in all the overlapping windows w_k that covers i, so the value of q_i in (4) is not identical when it is computed in different windows. A simple strategy is to average all the possible values of q_i. So after computing (a_k, b_k) for all windows w_k in the image, we compute the filtering output by :

$$q_i = \frac{1}{|w|}\sum_{k|i\in w_k}(a_k I_i + b_k) \tag{9}$$

Noticing that $\sum_{k|i\in w_k} a_k = \sum_{k\in w_i} a_k$ due to the symmetry of the box window, (9) is rewritten as :

$$q_i = \bar{a}_i I_i + \bar{b}_i \tag{10}$$

where $\bar{a}_i = \frac{1}{|w|}\sum_{k\in w_i} a_k$ and $\bar{b}_i = \frac{1}{|w|}\sum_{k\in w_i} b_k$ are the average coefficients of all windows overlapping i. The averaging strategy of overlapping windows is popular in image denoising.

With the modification in (10), ∇q is no longer scaling of ∇I because the linear coefficients $(\bar{a}_i, \bar{b}_i)$ vary spatially. But as $(\bar{a}_i, \bar{b}_i)$ are the output of a mean filter, their gradients can be expected to be much smaller than that of I near strong edges. In short, abrupt intensity changes in I can be mostly preserved in q.

Equations (7), (8), and (10) are the definition of the guided filter. A pseudocode is in Algorithm 1. In this algorithm, f_{mean} is a mean filter with a window radius r. The abbreviations of correlation (corr), variance (var), and covariance (cov) indicate the intuitive meaning of these variables.

Algorithm 1. Guided Filter.
Input: filtering input image p, guidance image I, radius r, regularization ϵ
Output: filtering output q.
1: $mean_I = f_{mean}(I)$
   $mean_p = f_{mean}(p)$
   $corr_I = f_{mean}(I.*I)$
   $corr_{Ip} = f_{mean}(I.*p)$
2: $var_I = corr_I - mean_I .* mean_I$
   $cov_{Ip} = corr_{Ip} - mean_I .* mean_p$
3: $a = cov_{Ip} ./(var_I + \epsilon)$
   $b = mean_p - a.* mean_I$
4: $mean_a = f_{mean}(a)$
   $mean_b = f_{mean}(b)$
5: $q = mean_a .* I + mean_b$

III. Overall Approach

The flowchart of the proposed image fusion method is shown in Fig. 3.1. We first employ guided filtering for the extraction of base layers and detail layers from the input images. qi computed in (9) preserves the strongest edges in I while smoothing small changes in intensity. Let $b_K(i',j')$ be the base layer computed from (9) (i.e.,$b_K(i',j') = qi$ and $1 \le K \le N$) for K^{th} input image denoted by $I_K(i',j')$. The detail layer is defined as the difference between the guided filter output and the input image, which is defined as

$$d_K(i', j') = I_K(i', j') - b_K(i', j') \quad (11)$$

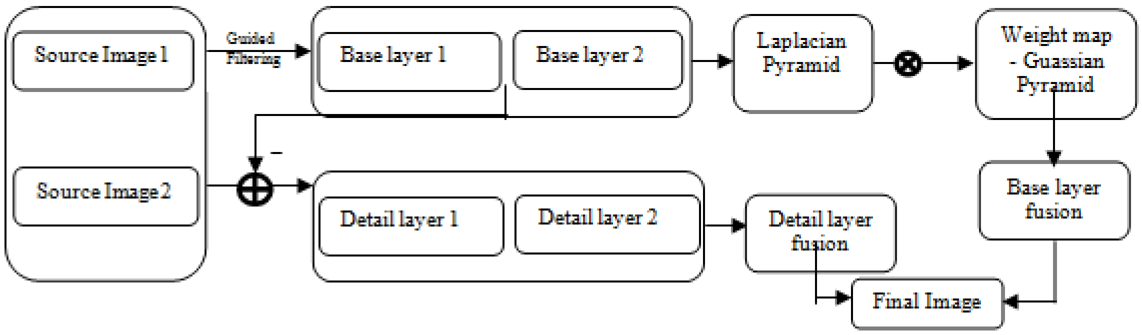

Fig 3.1. Flowchart of the proposed method

3.1 Base Layer Fusion

The pyramid representation expresses an image as a sum of spatially band-passed images while retaining local spatial information in each band. A pyramid is created by lowpass filtering an image $G0$ with a compact two-dimensional filter. The filtered image is then subsampled by removing every other pixel and every other row to obtain a reduced image $G1$.This process is repeated to form a Gaussian pyramid $G0$, $G1$, $G2$, $G3$, . . . , Gd. Expanding $G1$ to the same size as $G0$ and subtracting yields the band-passed image $L0$. A Laplacian pyramid $L0$, $L1$, $L2$, . . . , $Ld{-}1$, can be built containing band-passed images of decreasing size and spatial frequency.

$$L_l = G_l - G_{l+1}, \qquad l = 1, \ldots\ldots, d-1 \quad (12)$$

where l refers to the number of levels in the pyramid.

The original image can be reconstructed from the expanded band-pass images:

$$G0 = L0 + L1 + L2 + \cdots + Ld - 1 + Gd \quad (13)$$

The Gaussian pyramid contains low-passed versions of the original $G0$, at progressively lower spatial frequencies. This effect is clearly seen when the Gaussian pyramid levels are expanded to the same size as $G0$. The Laplacian pyramid consists of band-passed copies of $G0$. Each Laplacian level contains the edges of a certain size and spans approximately an octave in spatial frequency.

(a) Quality measures

Many images in the stack contain flat, colorless regions due to under- and overexposure. Such regions should receive less weight, while interesting areas containing bright colors and details should be preserved. The following measures are used to achieve this:

• Contrast: Contrast is created by the difference in luminance reflected from two adjacent surfaces. In other words, contrast is the difference in visual properties that makes an object distinguishable from other object and the background. In visual perception contrast is determined by the difference in color and brightness of the object with other object. It is the difference between the darker and the lighter pixel of the image, if it is big the image will have high contrast and in the other case the image will have low contrast.

A Laplacian filter is applied to the grayscale version of each image, and take the absolute value of the filter response. This yields a simple indicator C for contrast. It tends to assign a high weight to important elements such as edges and texture.

• Saturation: As a photograph undergoes a longer exposure, the resulting colors become desaturated and eventually clipped. The saturation of a color is determined by a combination of light intensity and how much it is distributed across the spectrum of different wavelengths. The purest (most saturated) color is achieved by using just one wavelength at a high intensity, such as in laser light. If the intensity drops, then as a result the saturation drops. Saturated colors are desirable and make the image look vivid. A saturation measure S is included which is computed as the standard deviation within the R, G and B channel, at each pixel.

• Exposure: Exposure is a term that refers to two aspects of photography – it is referring to how to control the lightness and the darkness of the image. In photography, exposure is the amount of light per unit area reaching a photographic film. A photograph may be described as overexposed when it has a loss of highlight detail, that is, when important bright parts of an image are washed out or effectively all white, known as blown out highlights or clipped whites. A photograph may be described as underexposed when it has a loss of shadow detail, that is, when important dark areas are muddy or indistinguishable from black, known as blocked up shadows. Looking at just the raw intensities within a channel, reveals how well a pixel is exposed. We want to

keep intensities that are not near zero (underexposed) or one (overexposed). We weight each intensity i based on how close it is to 0.5 using a Gauss curve:

$$exp\left(-\frac{(i-0.5)^2}{2\sigma^2}\right) \tag{14}$$

To account for multiple color channels, we apply the Gauss curve to each channel separately, and multiply the results, yielding the measure E.

The fused base layer $bf(i',j')$ is computed as the weighted sum of the base layers $b1(i',j'), b2(i',j'), \ldots, bN(i',j')$ obtained across N input exposures. Pyramidal approach is used to generate Laplacian pyramid of the base layers $L\{b_K(i',j')\}^l$ and Gaussian pyramid of weight map functions $G\{W_K(i',j')\}^l$ estimated from three quality measures (i.e., saturation $S_K(i',j')$, contrast $C_K(i',j')$, and exposure $E_K(i',j')$). Here, l ($0 < l < d$) refers to the number of levels in the pyramid and K ($1 < K < N$) refers to the number of input images. The weight map is computed as the product of these three quality metrics (i.e. $W_K(i',j') = S_K(i',j') \cdot C_K(i',j') \cdot E_K(i',j')$). The $L\{b_K(i',j')\}^l$ multiplied with the corresponding $G\{W_K(i',j')\}^l$ and summing over K yield modified Laplacian pyramid $L^l(i',j')$ as follows:

$$L^l(i',j') = \sum_{K=1}^{N} L\{b_K^l(i',j')\}\, G\{W_K^l(i',j')\} \tag{15}$$

The $b_f(i',j')$ that contains well exposed pixels is reconstructed by expanding each level and then summing all the levels of the Laplacian pyramid:

$$b_f(i',j') = \sum_{l=0}^{d} L^l(i',j') \tag{16}$$

3.2 Detail Layer Fusion

The detail layers computed in (11) across all the input exposures are linearly combined to produce fused detail layer $df(i',j')$ that yields combined texture information as follows:

$$d_f(i',j') = \frac{\sum_{K=0}^{N} \gamma(d_K(i',j'))}{N} \tag{17}$$

where γ is the user defined parameter to control amplification of texture details (typically set to 5).

Finally, the detail enhanced fused image $g(i',j')$ is easily computed by simply adding up the fused base layer $b_f(i',j')$ computed in (16) and the manipulated fused detail layer $df(i',j')$ in (17) as follows:

$$g(i',j') = d_f(i',j') + b_f(i',j') \tag{18}$$

3.3 Numerical Analysis

Numerical analysis is the process of evaluating a technique via some objective metrics. For this purpose, two fusion quality metrics [12], i.e., information theory based metric (Q_{MI}) [13] and structure based metrics (Qc) [14] are adopted. In order to assess the fusion performance, fusion quality metric is used.

(a) Normalized mutual information (Q_{MI})

Normalized mutual information, Q_{MI} is an information theory based metric. Mutual information improves image fusion quality assessments. One problem with traditional mutual information metric is that it is unstable and may bias the measure towards the source image with the highest entropy.

The size of the overlapping part of the images influences the mutual information measure in two ways. First of all, a decrease in overlap decreases the number of samples, which reduces the statistical power of the probability distribution estimation. Secondly, with increasing misregistration (which usually coincides with decreasing overlap) the mutual information measure may actually increase. This can occur when the relative areas of object and background even out and the sum of the marginal entropies increases, faster than the joint entropy. Normalized measure of mutual information is less sensitive to changes in overlap. Hossny *et al.* modified it to the normalized mutual information [13]. Here, Hossny *et al.*'s definition is adopted.

$$Q_{MI} = 2\left[\frac{MI(A,F)}{H(A)+H(F)} + \frac{MI(B,F)}{H(B)+H(F)}\right] \tag{19}$$

where $H(A)$, $H(B)$ and $H(F)$ are the marginal entropy of A, B and F, and $MI(A, F)$is the mutual information between the source image A and the fused image F.

$$MI(A,F) = H(A) + H(F) - H(A,F) \tag{20}$$

where $H(A, F)$ is the joint entropy between A and F, $H(A)$ and $H(F)$ are the marginal entropy of A and F, respectively, and $MI(B,F)$ is similar to $MI(A, F)$. The quality metric Q_{MI} measures how well the original information from source images is preserved in the fused image.

(b) Cvejic et al.'s metric (Qc)

Cvejic et al.'s metric, Qc is a structure based metric. It is calculated as follows :

$$Q_C = \mu(A_w, B_w.F_w)UIQI(A_w,F_w) + (1-\mu(A_w,B_w,F_w))UIQI(B_w,F_w) \quad (21)$$

where $\mu(A_w, B_w.F_w)$ is calculated as follows :

$$\mu(A_w,B_w,F_w) = \begin{cases} 0, & if \frac{\sigma_{AF}}{\sigma_{AF}+\sigma_{BF}} < 0 \\ \frac{\sigma_{AF}}{\sigma_{AF}+\sigma_{BF}}, & if\ 0 \le \frac{\sigma_{AF}}{\sigma_{AF}+\sigma_{BF}} < 1 \\ 1, & if \frac{\sigma_{AF}}{\sigma_{AF}+\sigma_{BF}} > 1 \end{cases} \quad (22)$$

σ_{AF} and σ_{BF} are the covariance between A,B and F, UIQI refers to the universal image quality index. The Qc quality metric estimates how well the important information in the source images is preserved in the fused image, while minimizing the amount of distortion that could interfere with interpretation.

IV. Results And Discussion

The system described above is implemented using Matlab and the result was successfully obtained. In this section, the obtained results are provided. Figure 4.1 shows the base and detail layers.

Figure 4.2: Base layers and Detail layers

Pyramidal approach is used for fusing base layers. Quality measures of images are considered to compute the weight map. Weight map is the combination of contrast, saturation and exposure. Figure 4.3 shows the gaussian pyramid of weight map function.

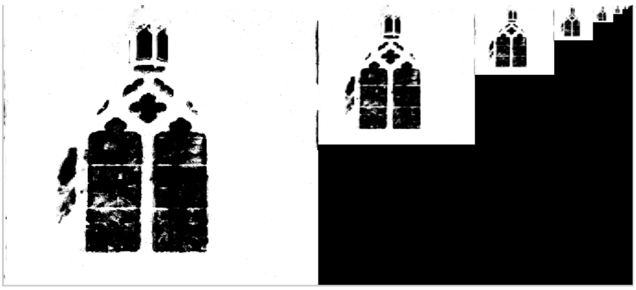

Figure 4.3: Gaussian pyramid

Laplacian pyramid of the base layers are generated. Thus obtained laplacian pyramid is shown in figure 4.4.

Figure 4.4: Laplacian pyramid

Fused pyramid is obtained by combining the Gaussian pyramid of weight map functions and Laplacian pyramid of base layers. Figure 4.5 shows the fused pyramid.

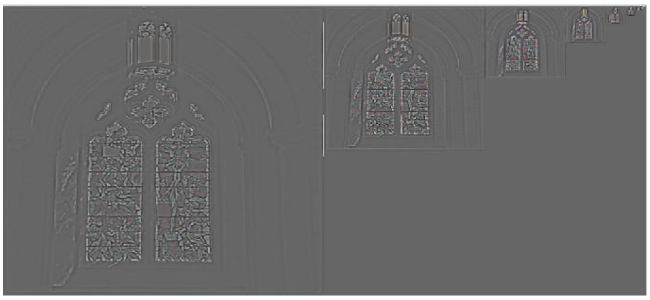

Figure 4.5: Fused pyramid

Fused base layer is the weighted sum of base layers. The detail layers obtained are boosted and fused. Figure 4.6 shows the fused base and detail layers.

Figure 4.6: Fused base layer and detail layer

Finally, the fused image is obtained by combining the obtained fused base and fused detail layers. The fused image is shown in figure 4.7. Numerical analysis is performed on the obtained results.

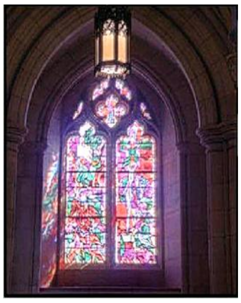

Figure 4.7: Fused image

The following shows the results obtained for some of the other source images.

(a) (b)

Figure 4.8: (a) Source Images (b) Fused Image

(a) (b)

Figure 4.9: (a) Source Images (b) Fused Image

(a) (b)

Figure 4.10: (a) Source Images (b) Fused Image

V. Conclusion

We proposed a technique for fusing multiexposure input images. The proposed method constructs a detail enhanced image from a set of multiexposure images by using a multiresolution decomposition technique. When compared with the existing techniques which use multiresolution and single resolution analysis for exposure fusion, the current proposed method performs better in terms of enhancement of texture details in the fused image. The framework is inspired by the edge preserving property of guided filter that has better response near strong edges. Experiments show that the proposed method can well preserve the original and complementary information of multiple input images.

REFERENCES

[1] S. Li, J. Kwok, I. Tsang, and Y. Wang, "Fusing images with different focuses using support vector machines," *IEEE Trans. Neural Netw.*, vol. 15, no. 6, pp. 1555–1561, Nov. 2004.

[2] G. Pajares and J. M. de la Cruz, "A wavelet-based image fusion tutorial," *Pattern Recognit.*, vol. 37, no. 9, pp. 1855–1872, Sep. 2004.

[3] D. Looney and D. Mandic, "Multiscale image fusion using complex extensions of EMD," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1626–1630, Apr. 2009.

[4] M. Kumar and S. Dass, "A total variation-based algorithm for pixellevel image fusion," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 2137–2143, Sep. 2009.

[5] P. Burt and E. Adelson, "The laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, Apr. 1983.

[6] O. Rockinger, "Image sequence fusion using a shift-invariant wavelet transform," in *Proc. Int. Conf. Image Process.*, vol. 3, Washington, DC, USA, Oct. 1997, pp. 288–291.

[7] J. Liang, Y. He, D. Liu, and X. Zeng, "Image fusion using higher order singular value decomposition," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2898–2909, May 2012.

[8] R. Shen, I. Cheng, J. Shi, and A. Basu, "Generalized random walks for fusion of multi-exposure images," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3634–3646, Dec. 2011.

[9] M. Xu, H. Chen, and P. Varshney, “An image fusion approach based on markov random fields,” *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 5116–5127, Dec. 2011.

[10] K. He, J. Sun, and X. Tang, “Guided image filtering,” in *Proc. Eur. Conf. Comput. Vis.*, Heraklion, Greece, Sep. 2010, pp. 1–14.

[11] N. Draper and H. Smith, *Applied Regression Analysis*. New York, USA: Wiley, 1981.

[12] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiere, and W. Wu,“Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 94–109, Jan. 2012.

[13] M. Hossny, S. Nahavandi, and D. Creighton, “Comments on ‘information measure for performance of image fusion’,” *Electron. Lett.*, vol. 44, no. 18, pp. 1066–1067, Aug. 2008.

[14] N. Cvejic, A. Loza, D. Bull, and N. Canagarajah, “A similarity metric for assessment of image fusion algorithms,” *Int. J. Signal Process.*, vol. 2, no. 3, pp. 178–182, Apr. 2005.